Predictive Validity of *DSM-III* and Feighner Definitions of Schizophrenia

A Comparison With Research Diagnostic Criteria and CATEGO

John E. Helzer, MD; Ian F. Brockington, MD; R. E. Kendell, MD

• The new *DSM-III* and the Feighner criteria give definitions of schizophrenia that are restrictive in that they require six months of illness before the diagnosis can be made. The definitions are compared with definitions of schizophrenia in the Research Diagnostic Criteria and CATEGO in terms of predictive validity using a cohort of 125 psychotic patients from the United States-United Kindgom Diagnostic Project. The results showed that while the *DSM-III* and Feighner definitions were more restrictive, they also had a higher degree of predictive specificity than the other two definitions. No patient diagnosed by *DSM-III* or Feighner criteria showed a change in diagnosis over time, and these patients were more likely to have presistent schizophrenic symptoms and to have social handicap. The appropriateness of the requirement of six months of illness and the practicality of using outcome as a validity measure are discussed.

(*Arch Gen Psychiatry* 1981;38:791-797)

One of the most useful aspects of diagnosis is prediction of outcome or prognosis, but discovering a definition of schizophrenia that predicts uniform outcome has been particularly difficult. In an extensive review, Pope and Lipinski found little evidence that schizophrenic symptoms cross-sectionally allowed any uniform prediction of outcome.[1] Similarly, Carpenter et al, using International Pilot Study of Schizophrenia (IPSS) data, have found that cross-sectional schizophrenic symptoms are only weakly related to outcome.[2]

One clinical variable that appears to be useful in predicting outcome in schizophrenia is mode of onset. Sudden onset is almost always cited as one of the indicators of good outcome in studies that attempt to differentiate good and poor-prognosis schizophrenia.[3-5] Even in schizoaffective states, mode of onset is an effective predictor.[6] Two taxonomies used in the United States, the recently published *DSM-III*[7] and the Feighner criteria,[8] require an illness of six months' duration before a diagnosis of schizophrenia can be given. Bland and Orn have suggested that the Feighner criteria may be more successful than definitions not requiring six months of illness in predicting a poor outcome in schizophrenia.[9]

In a recent follow-up study of psychotic patients,[10] it was shown that some definitions of schizophrenia (the Research Diagnostic Criteria [RDC],[11] CATEGO,[12] Langfeldt's,[13] and Carpenter's flexible system[14]) were better predictors of outcome than others (eg, Schneider's first-rank symptoms[15]). However, even the better definitions were not especially good in predicting a poor outcome. For example, of the 44 patients diagnosed schizophrenic by the RDC,[11] 20 (45%) did not have schizophrenic phenomena during the follow-up period. The Feighner definition[8] was not examined in this earlier report because it appeared too restrictive in another diagnostic exercise, selecting only ten patients as schizophrenic from 119 psychotic patients at first admission.[10] However, while *DSM-III* and Feighner criteria can be expected to be restrictive because of the requirement of six months of illness, their specificity in terms of predicting poor outcome may be higher than other systems'.

The purpose of the present study is to examine the predictive specificity of the definitions of schizophrenia and schizophreniform disorder in *DSM-III* and of the Feighner criteria for schizophrenia using the same series of patients that was used for the comparison previously noted.[10] We have contrasted these two definitions with the RDC and an inclusive CATEGO definition that appeared to be fairly successful in the earlier comparison.

PATIENTS AND METHODS

Case Examination

Case selection, interviewing, and follow-up procedures have been detailed previously.[10] To review briefly, 134 psychotic patients were interviewed at Netherne Hospital near London in 1966-1968 by Cooper, Kendell, and Sartorius as part of the United States-United Kingdom (US/UK) Diagnostic Project.[16] Interviews were done using the seventh edition of the Present State Examination[12] and a psychiatric history schedule. These patients were traced and reexamined by one of us (I.F.B.) in 1972-1974, an average of 6.5 years after the index admission. From the original sample, six patients were excluded because of lack of follow-up information and three were excluded because at the time of the present review the mental state schedule was missing. This left 125 patients for the present study.

Diagnosis

In the diagnostic analysis previously reported,[10] Kendell and Leff made diagnoses of schizophrenia using the RDC[11] and Langfeldt's,[13] Carpenter's,[14] Schneider's,[15] and Astrachan's[17] criteria. Diagnoses made by the physicians of Netherne Hospital at index admission, CATEGO diagnoses,[12] and US/UK project diag-

---

Accepted for publication Jan 22, 1981.

From the Department of Psychiatry, Washington University School of Medicine, St Louis (Dr Helzer); the Institute of Psychiatry, London (Dr Brockington); and the Royal Edinburgh Hospital, Edinburgh (Dr Kendell).

Reprint requests to Department of Psychiatry, Washington University School of Medicine, 4940 Audubon Ave, St Louis, MO 63110 (Dr Helzer).

noses were also available. For the present study, two authors reviewed the original material again in order to apply Feighner[8] and RDC[11] definitions of schizophrenia and *DSM-III*[7] definitions of schizophrenia and schizophreniform illness. We independently studied history and mental state schedules, taking about 15 minutes per case, and gave each patient a diagnosis in each of the three systems. We then compared notes and resolved differences, reaching an agreed diagnosis for each system.

There were methodological difficulties in that the interview schedules were designed for a different diagnostic scheme than the ones we were using. Thus, ratings needed for particular criteria were often not present in the form needed. For some criterion items there was a wealth of information. For example, in the assessment of premorbid adjustment, both patient and informant ratings were available for educational and occupational history, sexual and marital history, and previous personality and friendships. Conversely, some items lacked necessary information. For example, schizophrenic illness in family members (required for the Feighner definition) was rated in the schedules for parents, siblings, and offspring, but there were no ratings for second-degree relatives. Other items were an approximation of what we required. For example, a six-month duration was embedded within a code that encompassed illness lasting from three to 12 months. Whenever possible, we relied on marginal notes for clarification when such problems arose. The most difficult part of the task was that there were several interview items relevant to each criterion item, and these were scattered throughout the schedules.

Outcome Measures

At the time of the follow-up interview, 86 ratings were made about psychopathologic symptoms, social state, and pattern of illness during the follow-up period. A principal components analysis of the psychopathologic symptom ratings was used to select nine parameters of symptoms that were then rated directly on visual analogue scales.

Data on all subsequent patients admitted to the mental hospital were transcribed and the patients were diagnosed by R.E.K. without knowing the identity of the patient. All follow-up data (including data from each subsequent admission and the follow-up interview) were transcribed into a case summary of about 1,000 words. These summaries were studied and the patients were diagnosed independently by I.F.B. and R.E.K., leading to agreed follow-up diagnoses based on the follow-up period alone. The follow-up diagnoses were then compared with the original diagnoses, and a final diagnosis was made based on all sources of information.

From all these ratings, scores, and diagnoses, the following items were selected for this report:

1. From the original outcome ratings, we determined the "definite presence" at some time during the follow-up period of manic symptoms, delusions of control, hallucinations, first-rank symptoms of Schneider, apathy, flattening of affect, and thought disorder, and also the "persistent presence" of depressed mood and somatic symptoms of depression.
2. Using the blind diagnoses of patients subsequently admitted, we noted which patients had at least one episode diagnosed as schizophrenia, at least one of mania, and at least one of depression.
3. We took the mean of the two visual analogue scales measuring employment record and social involvement and used this as a "combined social status" rating.
4. We noted the scores on the following seven scales measuring the severity of dimensions of psychopathologic symptoms: depression, manic symptoms, passivity experiences, auditory hallucinations, delusion formation, persecution, and defect state.
5. We used the final diagnoses made by I.F.B. and R.E.K.
6. We derived two summary scores for each patient as follows.

The "outcome regression score" was obtained by dividing the patients into five groups representing five grades of outcome from worst to best using five variables concerned with overall outcome, irrespective of symptoms, to predict this grouping. The five variables and their weights were as follows: employment record, +0.66; social involvement, +0.65; clinical recovery, +0.53; percentage time in the hospital, +0.45; and number of episodes, −0.11. The "psychopathology summary score" was obtained by comparing two criterion groups typical of schizophrenia and affective psychosis with respect to the seven scales mentioned in item 4 using discriminant function analysis to derive weights best predicting membership in schizophrenic and affective groups. The variables and their weights were as follows: auditory hallucinosis, −0.57; delusion formation, −0.52; defect state, −0.49; passivity experiences, −0.28; persecution, −0.18; somatic depressive symptoms, +0.45; and manic symptoms, +0.54. A high "outcome regression score" was indicative of poor outcome, and a negative psychopathology summary score indicated that the balance of psychopathologic symptoms were schizophrenic.

Table 1.—Agreement Between Examiners After Independent Case Review

Diagnosis	Concordance, κ I.F.B. and J.E.H.	Concordance, κ I.F.B./J.E.H. and R.E.K./J.P.L.*
Feighner criteria		
Probable schizophrenia	.38	. . .
Definite schizophrenia	.50	. . .
Undiagnosed	.52	. . .
DSM-III criteria		
Schizophreniform disorder	.00	. . .
Schizophrenia	.49	. . .
Research Diagnostic Criteria		
Probable schizophrenia	.21	.70
Definite schizophrenia	.45	.68

*J.P.L. is J. P. Leff, MD.

Groups Compared

For the purpose of this comparison, we combined Feighner definite and probable schizophrenia, RDC definite and probable schizophrenia, and all of the schizophrenic groupings in CATEGO (S+, P+, O+, S?, P?, and O?). We chose the most inclusive CATEGO definition because previous work suggested that this was more successful in predicting long-term prognosis than using either the S+ or S+, P+, O+ groupings.[10,18] This gave us four groups of schizophrenics (*DSM-III*, Feighner, RDC, and CATEGO) that we compared individually with all of the other psychotic patients in the series. In addition, we made specific comparisons of (1) *DSM-III* schizophrenia with *DSM-III* schizophreniform disorder, (2) RDC schizophrenics who also fulfilled *DSM-III* criteria for schizophrenia with those who did not, and (3) CATEGO schizophrenics who did and did not fulfill *DSM-III* criteria.

RESULTS

Concordance Between Examiners

These results are given in Table 1. It was necessary to discuss some aspect of the diagnosis in 91 of the 125 cases. Patients were diagnosed independently in batches of from ten to 34 patients, and diagnoses were compared after each batch. All ten of the first batch had to be discussed, as did 25 of 32 in the second batch, then 14 of 24, 16 of 25, and 26 of 34 in subsequent batches. It can be seen that there was little tendency for total agreement to improve with time; however, there were few major disagreements. Most of the discussions were of minor differences in the interpretation of a rating or a hand-written note or were based on a bit of information that one of us found but that the other had overlooked.

We also compared figures for the agreed diagnosis of RDC schizophrenia made by J.E.H. and I.F.B. with those in the same series of patients using the same material made by R.E.K. and J. P. Leff, MD, for an earlier investigation.[10] The latter two diagnosed definite schizophrenia in 33 of the 125 patients and probable schizophrenia in eight others. Twenty-three definite schizophrenics and five probable schizophrenics were diagnosed by J.E.H. and I.F.B. The coefficients of agreement (κ) between the two pairs of raters are .68 for definite schizophrenia and .70 if probable schizophrenia is included (Table 1).

Table 2.—Concordance (κ Value) Between Definitions of Schizophrenia*

Definition	N	DSM-III	RDC	CATEGO (All)	Langfeldt	US/UK Diagnostic Project
Feighner, definite or probable	18	.84	.68	.24	.45	.43
DSM-III	19	...	.71	.26	.48	.46
RDC, definite or probable	28	...	...	.44	.67	.58
CATEGO, all schizophrenic groupings	52	...	...	...	.62	.50
Langfeldt	36	...	...	...	...	.63

*RDC indicates Research Diagnostic Criteria; UK, United Kingdom.

Table 3.—General Outcome Data

Diagnosis	N	No. of Cases With Full Recovery	% of Follow-up Period Spent in Hospital	Combined Social Status Score	Outcome Regression Score
DSM-III schizophrenia	19	7*	45†	54.0*	2.53*
Feighner criteria					
Definite schizophrenia only	13	6*	49*	52.6*	2.49*
Definite or probable schizophrenia	18				
Research Diagnostic Criteria					
Definite schizophrenia only	23	12*	34	45.9†	2.11†
Definite or probable schizophrenia	28				
CATEGO, all schizophrenic groupings	52	30†	26	42.9‡	1.88‡
Whole series	125	89	24	38.1	1.67

*Different from the whole series at $P < .001$.
†Different from the whole series at $P < .01$.
‡Different from the whole series at $P < .05$.

Concordance Between Definitions

Eighteen patients met the Feighner criteria for definite or probable schizophrenia, and 40 were classified as having "undiagnosed psychiatric illness" by that system. These latter patients were usually patients who appeared schizophrenic cross-sectionally but had an insufficient duration of illness to meet criteria or were patients with a mixture of schizophrenic and affective symptoms. Nineteen patients received a *DSM-III* diagnosis of schizophrenia and seven received a diagnosis of schizophreniform psychosis, while nine were undiagnosed. Fewer patients were diagnosed as schizophrenic using these two systems than with the RDC (28 definite or probable cases), the original US/UK Diagnostic Project team (46 cases), CATEGO (35 S and 17 other schizophrenic categories), Schneider's criteria (36 cases), Langfeldt's criteria (36 cases), or the Netherne Hospital physicians (61 cases).

As given in Table 2, concordance between the *DSM-III* and Feighner definitions was high ($\kappa = .84$). Concordance was low between these definitions and CATEGO, intermediate with the US/UK Project diagnosis and Langfeldt's criteria, and fairly high with the RDC. Concordance between CATEGO, Langfeldt's criteria, and US/UK Diagnostic Project diagnosis were fairly high. Thus, there seemed to be the following two groupings: *DSM-III* and Feighner criteria on the one hand and CATEGO, Langfeldt's criteria, and the US/UK Diagnostic Project diagnoses on the other, with relatively high mutual concordances but low concordance between the two groups. The RDC showed fairly good concordance with both groupings.

Outcome

Table 3 compares patients meeting the four definitions of schizophrenia with respect to the degree of recovery from the index episode, hospitalization during the follow-up, social status at follow-up, and the general outcome regression score. The results for CATEGO and RDC are the same as those already published,[10] though the series is slightly smaller because of the omission of three patients whose schedules could not be found. The RDC criteria are rather more successful than CATEGO whichever measure is used, but they are both fairly successful in identifying a group of patients with relatively poor outcome. However, *DSM-III* and Feighner criteria are more successful than either the RDC or CATEGO.

Indices of Mania and Depression

The results are given in Table 4. None of the patients with *DSM-III* or Feighner schizophrenia was given a diagnosis of mania in any subsequent admission, and few of these patients showed even a hint of manic symptoms during the follow-up period. Similarly, few showed any notable degree of depression, and none received that final diagnosis. The RDC were about equally successful in eliminating those who showed marked depressive symptoms but not quite as successful in eliminating those with manic symptoms. The CATEGO criteria were relatively unsuccessful in eliminating manic patients and less successful in eliminating depressed patients; indeed, ten (19%) of the 52 patients with a CATEGO diagnosis of schizophrenia received a final diagnosis of affective disorder.

Schizophrenic Symptoms

Both *DSM-III* and Feighner schizophrenics showed much evidence of schizophrenic symptoms during the follow-up period (Table 5). When they are compared with RDC schizophrenics, there are few differences whether psychopathologic symptoms are measured by their definite presence at some time during the follow-up period or, as given in the Table, by their severity as rated on the scales. The ratings of defect symptoms are more severe in patients meeting *DSM-III* and Feighner criteria; the summary score, which takes account of the balance of affective and schizophrenic symptoms, showed higher schizophrenic scores than in the RDC schizophrenics. The CATEGO schizophrenics generally showed somewhat fewer schizophrenic pathologic symptoms and the psychopathology summary score showed that the balance of symptoms during the follow-up was affective. Only 16 (31%) of the 52 CATEGO schizophrenics had subsequent admissions that were blindly diagnosed as schizophrenia during the 6.5-year follow-up period. However, this did not discriminate much between the definitions since only about 50% of the *DSM-III* and

Table 4.—Indices of Mania and Depression

		Indices of Mania			Indices of Depression			
		No. of Patients			No. of Patients			
	N	Probably Manic at Some Time	Definitely Manic at Some Time	Manic Scale Score	Depression Present for Long Periods	Somatic Symptoms of Depression Present for Long Periods	Depression Scale Score	Final Diagnosis of Affective Psychosis (Unipolar or Bipolar)
DSM-III schizophrenia	19	3*	1	3.4*	1*	0	11.8*	0†
Feighner, definite or probable schizophrenia	18	4	1	3.6*	1	0	11.6*	0†
Research Diagnostic Criteria, definite or probable schizophrenia	28	7	4	7.0	2*	0*	12.1‡	1†
CATEGO, all schizophrenic groupings	52	19	9	11.0	8*	2‡	17.6*	10†
Whole series	125	53	26	13.7	34	18	23.2	66

*Different from the whole series at $P < .05$.
†Different from the whole series at $P < .001$.
‡Different from the whole series at $P < .01$.

Table 5.—Indices of Schizophrenic Symptoms

		Scaled Scores					General Indices		
Criteria	N	Passivity	Auditory Hallucinations	Delusions	Persecution	Defect Scale	≥1 Admission Diagnosed Schizophrenia	Final Diagnosis of Schizophrenia	Psychopathology Summary Score
DSM-III	19	18.0*	15.8†	38.7*	26.1†	26.1*	9‡	18*	−0.51*
Feighner	18	18.8*	16.7†	39.6*	26.9†	24.9*	9‡	18*	−0.53*
Research Diagnostic Criteria	28	18.7*	14.1†	36.6*	25.3‡	18.8‡	12*	25*	−0.25*
CATEGO	52	12.4*	12.7‡	29.1*	21.0†	13.1†	16†	37*	+0.29*
Whole series	125	5.9	7.7	20.1	16.0	8.8	24	50	+0.87

*Different from the whole series at $P < .001$.
†Different from the whole series at $P < .05$.
‡Different from the whole series at $P < .01$.

Table 6.—Comparison of *DSM-III* Schizophrenia and Schizophreniform Disorder

		Pathologic Symptom Scale Scores						General Outcome Measures		
	N	Manic	Passivity	Auditory Hallucination	Delusion	Persecution	Defect	% of Time in Hospital	Combined Social Status Score	Outcome Regression Score
Schizophrenia	19	3.4	18.0	15.8	38.7	26.0	26.1	46	54	2.53
Schizophreniform disorder	7	13.0	20.4	13.4	34.0	20.9	8.1	12	34	1.46
P	. . .	.04	NS	NS	NS	NS	NS	.02	.03	.02

Feighner patients had subsequent admissions diagnosed as schizophrenia.

Comparison of *DSM-III* Schizophrenia and Schizophreniform Disorder

The 19 *DSM-III* schizophrenics were compared with seven patients with the same symptoms but a history of at least two weeks, but less than six months, of symptoms (schizophreniform disorder) (Table 6). There are few differences in the scales of schizophrenic pathologic symptoms as measured at the follow-up interview. There was a notable difference in the defect scale, but there were too few schizophreniform patients for this to be statistically significant. Schizophreniform patients did show significantly more manic symptoms. Despite similarities in schizophrenic pathologic symptoms occurring during the follow-up, there were significant differences in three measures of general outcome, with schizophreniform patients having a much better outcome.

Comparison of RDC and CATEGO Schizophrenics Who Did and Did Not Meet *DSM-III* Criteria

Ten of the 28 RDC schizophrenics did not meet *DSM-III* criteria. The differences between the two groups are similar to those between *DSM-III* schizophrenia and schizophreniform disorder. None of the differences in psychopathologic symptoms occurring during follow-up was significant, but major differences again were found in the general outcome measures, where *DSM-III* schizophrenics were worse in every respect.

A comparison of CATEGO schizophrenics meeting *DSM-III* criteria with those who did not shows large differences not only in general outcome measures but also in psychopathologic symptoms. There was a particularly large difference in the percent of time spent in the hospital during the follow-up period (53% vs 17%, $P < .0001$). All of the CATEGO schizophrenics who met the *DSM-III* definition were given a final diagnosis of schizophrenia, whereas this was the case for only 61% of those who did not meet the *DSM-III* definition.

METHODOLOGICAL DISCUSSION

Limitations of the Study

The Netherne series of psychotic patients was not ideal for a comparison of schizophrenic definitions. These were not necessarily first admissions, and some of the patients (especially the 12 with a US/UK Diagnostic Project diagnosis of residual schizophrenia) had no florid symptoms at the time of their index admission. They would not be diagnosed as schizophrenic by CATEGO, which relies on the presence of specific and florid symptoms at the time of examination. The series is small, and only 68 patients satisfied any definition of schizophrenia.

The data are extremely good in other respects. The mental state examinations were executed thoroughly by experts, and the schedule was very detailed, with over 400 defined ratings. The history and other schedules were also quite detailed, and in many cases the historical information had been checked with next of kin.

Though a wealth of information was available for each case, we did have some difficulty in assimilating the data. Information relevant to the diagnoses we rated was scattered, and there was less description and direct reporting of patients' statements than we would have liked. For patients not currently ill, we were often dependent on marginal notes for information regarding specific symptoms. We did not have hospital physician notes, nursing records, or social work reports, and a summary of these types of data for both index and previous admissions would have been helpful.

Dr Brockington was not blind to outcome, having met and personally examined almost all of the patients five to eight years after index admission. The potential bias introduced by this was modified by the fact that J.E.H. was blind to all follow-up material after the index admission and remained so until after the completion of the study. Any disagreements between the independent diagnoses of these two raters were resolved entirely on the basis of the index ratings. We also reexamined several of the outcome measures using J.E.H.'s diagnoses only. On the whole, we found these to be just as predictive as our joint diagnoses. Both raters were exposed to some longitudinal data for those patients who had been admitted previously. However, such cases could meet criteria for schizophrenia in any of the three classifications that we rated and so should not favor one over the others in terms of predicting outcome.

Concordance Measures

The interrater concordances we report for applying the definitions of schizophrenia are low (none higher than a κ value of .52). This suggests that it is desirable to use at least two raters in diagnostic work of this kind. We believe the concordances reported here are representative of those that can be achieved when complex diagnostic rules are applied by experienced raters to detailed material. It has been shown previously that diagnostic concordance between raters using case summaries is generally lower than that obtained using live interviews.[19] It should be borne in mind that the diagnostic task was made even more complex by the fact that the schedules we rated were not originally designed for the diagnostic definitions we used. For these reasons, we used consensus diagnoses since these are more reliable than the diagnostic judgments of a single rater.[20] Reliabilities of the definitions of schizophrenia used here have been examined previously under more appropriate conditions.[12,21,22]

Concordance between definitions (Table 2) is not a measure of validity but is of interest in its own right and provides a better understanding of the reasons for differences in outcome in the various groups.

Predictive Validity of the Definitions Studied

In spite of the limitations, the data are sufficient to demonstrate differences in outcome between quite small groups of patients. There were no important outcome differences between 18 with probable or definite Feighner schizophrenia and the 19 with *DSM-III* schizophrenia, and in fact they proved to be highly concordant groups ($\kappa = .84$). The psychopathologic symptom criteria of these two definitions differ considerably. The Feighner definition emphasizes family and social variables more and is far less specific about the types of psychopathologic symptoms that are required. What these two definitions share is the stipulation that the disorder must have lasted at least six months. An interesting parallel comparison is between *DSM-III* definitions for schizophrenia and schizophreniform disorder (Table 6). Here, there is a clear difference in outcome that is demonstrable in the series of only 26 patients. The psychopathologic symptom criteria for these two diagnoses are identical, and the only difference is the requirement of six months of illness. It is clear that this latter criterion is the main predictor of the measures we have included under general outcome, eg, time in hospital and social status, as well as the absence of manic symptoms (Table 6).

In the comparisons of patients meeting RDC or CATEGO definitions but not meeting the *DSM-III* definitions, there are fewer significant differences in psychopathologic symptoms occurring during the follow-up period but consistent differences in the general outcome scales. This suggests that patients meeting any of these definitions are prone to recurrences of psychotic symptoms but that the *DSM-III* (and Feighner) definitions are strongly predictive of a degree of chronicity of illness associated with social decline.

The final diagnosis of schizophrenia was based on a review of all available material about the index admission and follow-up period and was done by two raters (R.E.K. and I.F.B.) who applied their own clinical criteria in arriving at an overall diagnostic judgment. Against this validity standard[23,24] the *DSM-III* and Feighner definitions performed well. All of the Feighner and 95% of the *DSM-III* schizophrenics had a final diagnosis of schizophrenia, while 89% of the RDC and only 71% of the CATEGO schizophrenics did.

COMMENT

Outcome as a Validity Measure

Questions can be raised about both the appropriateness and practicality of outcome as a measure of validity. The choice of a validating measure is to some degree arbitrary and is obviously a function of the goals of the taxonomist. But one use of diagnosis is to assist in predicting outcome both for the purpose of prognostication and as a basis for the systematic evaluation of intervention efforts.

Another use of a classificatory system is the assistance this provides in the discovery of etiology. Any measure of validity helps refine the classification system and make it more homogeneous. Such refinements directly benefit our ability to communicate and help in conceptualizing the scope of a particular illness, both of which directly enhance etiological research. Whether outcome is the best validating criterion for definitions that will direct etiological studies is questionable since it is clear that factors that are unique to the individual account for a great deal of the variation in outcome, even in illness states where etiology is known. But, as Robins and Guze have said, "The same illness may have a variable prognosis, but until we know more about the fundamental nature of the common psychiatric illnesses marked differences in outcome should be regarded as a challenge to the validity of the original diagnosis."[25]

One problem in quantifying predictive validity is that outcome is not a unitary trait. It is hard to know what the most appropriate measures of outcome are. The measurement of whatever outcome functions are chosen is likely to be difficult, and the correlation between various measures of outcome may be poor.[26] In the present study, we have reported outcome in terms of persisting psychopathologic symptoms, social function (employment and social involvement), continued hospitalization, and change in diagnosis over time.[25-27]

To our knowledge, this is the first outcome study of the *DSM-III* definition of schizophrenia, but there are previous outcome studies of the Feighner definition. Winokur and colleagues applied these criteria in a chart review of patients admitted to a state hospital in the 1930s and 1940s and then did a personal follow-up.[28,29] Only 8% of the patients they diagnosed as schizophrenic using the Feighner criteria had recovered during the follow-up period. The rate of recovery in schizophreniform psychosis (ie, patients with a chart diagnosis of schizophrenia who failed to meet the Feighner definition) was almost three times as great. Tsuang and colleagues have recently reported a 34-year follow-up of these cases and are still able to show large differences in outcome between schizophrenia and other diagnostic groups even after controlling for the potential selection-criteria biases in marital and occupational status.[30,31]

These studies, like ours, were of a cross section of hospitalized patients, not necessarily first admissions. Bland and Orn[9] reported a 14-year follow-up of first-admission cases in which they compared the following three sets of criteria for schizophrenia: Schneider's,[15] Astrachan's New Haven Schizophrenia Index,[17] and the Feighner criteria.[8] They found that only the Feighner diagnosis predicted outcome and that a rating scale composed of the nine criterion items explained 24% of the outcome variance. They did not use the Feighner criteria exactly as specified (ie, the requirement of six months of illness was not absolute but was used to differentiate probable from definite diagnoses). It would have been interesting to see if the criteria were even more powerful in predicting outcome if this requirement had been absolute.

Numerous previous investigators have demonstrated the poor predictability of schizophrenic symptoms. In a recent comprehensive review, Pope and Lipinski found that "Schizophrenic symptoms, taken alone and in cross section, have remarkably little, if any, demonstrated validity in determining diagnosis, prognosis, or treatment response in psychosis."[1] They concluded that there are no pathognomonic schizophrenic symptons. A previous study using US/UK data from Netherne found that while florid symptoms of psychosis can be reliably recognized by independent examiners, they predict outcome poorly.[10] Strauss and Carpenter[26] also compared various definitions of schizophrenia and found that, used cross-sectionally, narrow definitions did not predict outcome any better than broad ones. They were unable to find any set of cross-sectional criteria that predicted poor outcome at five years. Thus, the conclusion would seem to be that psychotic patients, no matter how floridly ill at one point in time, do not necessarily have a chronic illness, and either florid psychopathologic symptoms or the possession of particular symptoms without reference to duration of symptoms is an inadequate basis for prognosis. In their comparison of schizophrenics from developed and developing countries using IPSS data, Sartorius et al call for the consideration of social and marital factors and mode of onset for improving the prediction of prognosis in schizophrenia.[32] The data reported here clearly support that idea and suggest that mode of onset is the most powerful factor of the three.

Strauss and Carpenter have criticized the use of duration of schizophrenic symptoms as a means of improving predictive validity. They contend that "Since established chronicity predicts chronicity for most psychiatric disorders, its addition as a diagnostic criterion also negates the utility of poor outcome as a validating criterion for the existence of a specific schizophrenic process."[33] However, the six-month requirement must be taken in context. First, what is required is not six months of florid symptoms but rather some degree of illness. The *DSM-III* specifically allows inclusion of prodromal and residual symptoms in the six months. Second, this criterion permits a more accurate prediction of outcome over an average of 6.5 years in the present study, 14 years in the Bland and Orn study,[9] and 30 to 40 years in the Tsuang et al report.[30] In contrast to the up to 40 years of outcome being predicted, six months of illness is relatively brief and can hardly be seen as established chronicity. Finally, the six-month requirement is arbitrary and may be longer than necessary. Tsuang et al are currently investigating this issue in a sample of 315 patients in their 1940s cohort who were diagnosed as schizophrenic but who did not meet Feighner criteria because of, for some patients, a short duration of symptoms.[30] Cohorts like theirs will be of help in specifying empirically the duration of illness necessary before a schizophrenic outcome can be predicted with confidence. Some attention to mode of onset seems necessary to bring some order to the chaotic range of outcomes predicted by schizophrenic symptoms alone.

Narrow and Broad Definitions

The Feighner and *DSM-III* definitions are very narrow and identify only a small proportion of cases as schizophrenic. The appropriateness of this for predicting outcome finds considerable support but may be challenged on other grounds. For one thing, the effect of restrictive categories on the diagnostic system as a whole must be considered. It may be easier to produce homogeneity of outcome with a narrow definition of any diagnostic concept, but as Spitzer and Williams have observed in a discussion of the "coverage" of taxonomies, a classification of mental disorder that includes only narrowly defined, well-validated categories will not have specific categories for a sizable group of psychiatrically ill persons.[34] It is not possible to have narrow definitions of all diagnostic terms simultaneously without having a large residual category in the system. This is not necessarily undersirable, since the use of residual categories may permit greater confidence in the validity of the discrete categories and since the eventual goal is a series of discrete categories covering all illness. But the effect on the entire system of a narrowly defined portion must be recognized.

Another possible challenge to narrowly defined schizophrenia comes from a study by Shields and Gottesman of a spectrum of criteria from very narrow to very inclusive. They found that a "middle of the road" diagnosis maximized the monozygotic-dizygotic discrimination in a series of 114 twins.[35] Thus, on genetic grounds, the *DSM-III* and Feighner definitions might be too strict. However, there is no single correct measure of validity, and the appropriateness of the validity measure depends primarily on the reason for formulating the definition in the first place. If narrow definitions are useful for one important function (eg, prediction of outcome), it would not seem appropriate that they be discarded because they are not as useful for another important function (eg, genetic discrimination). An alternative approach is to develop discrete definitions of illness states and then combine these in various ways to maximize utility for a given usage or even to coalesce discrete groups if they are shown to function the same way across all validity measures. This process could help elucidate etiology. Taking the previous example, it would appear that a narrow definition of schizophrenia maximizes prediction of outcome, but a broader definition may maximize genetic discrimination. This might suggest that we look for environmental factors that predict a poor-outcome phenotype within the broader schizophrenic genotype.

Definitions of illness obviously can influence results, and here again the use of discrete definitions that can be disaggregated or coalesced would seem to have the best chance of maximizing generalizability while at the same time maximizing the opportunity for revealing important interactions. An example of this is the report by Sartorius et al from the IPSS that schizophrenics in developing countries have a better prognosis than those in developed countries.[32] The IPSS did not use a duration-of-illness variable in the definition of schizophrenia. Twenty-eight percent of the identified patients had a length of episode of inclusion of less than one month, and another 28%, one to three months.[32] Since sudden onset predicts good outcome in functional psychoses,[3-5] it is conceivable that the IPSS findings might be explained on the basis of ascertainment. A chronic schizophrenic illness, or one with an insidious onset, is more compatible with remaining in the community and even low level-employment than is a sudden-onset psychosis. If, as one would expect, mental health services are less accessible in developing countries, it might be more likely that a bias toward ascertainment of acute psychoses exists in developing countries, whereas in developed countries, which have more accessible mental health services, treatment is more likely to be obtained for chronic or insidious as well as sudden-onset illness.

For many diagnosticians in the United States, the *DSM-III* will mean the use of a definition of schizophrenia that is considerably more restrictive than they are accustomed to. In fact, the pendulum seems to have swung the other way from the time of the US/UK diagnostic study from which our data have been derived. At that time, the concept of schizophrenia in the United States was much broader than that in England and, presumably, Europe generally. Now the definition adopted by the American Psychiatric Association in *DSM-III* appears more restrictive than the current European concept. But the major advantage of the *DSM-III* concept is that it is also very specific. In terms of definition it is specific because the exact requirements for the diagnosis are enumerated. But perhaps more importantly, it is also very specific in the prediction of outcome, predicting with a high degree of accuracy patients who will remain chronically ill with a psychotic illness (ie, patients who, according to this restrictive definition, will remain schizophrenic).

This work was supported in part by Public Health Service grants 278-79-0017DB and MH-33883 and, during the follow-up and data collection, the Research Fund of the Bethlem Royal and Maudsley Hospitals, London (Dr Brockington).

Irving I. Gottesman, PhD, Gregory Carey, PhD, and Janet Williams, MSW, read and commented on the manuscript.

References

1. Pope HG Jr, Lipinski JF Jr: Diagnosis in schizophrenia and manic-depressive illness: A reassessment of the specificity of 'schizophrenic' symptoms in the light of current research. *Arch Gen Psychiatry* 1978;35:811-828.
2. Carpenter WT, Bartko JJ, Strauss JS, et al: Signs and symptoms as predictors of outcome: A report from the International Pilot Study of Schizophrenia. *Am J Psychiatry* 1978;135:940-945.
3. Stephens JH, Astrup AC: Prognosis in 'process' and 'nonprocess' schizophrenia. *Am J Psychiatry* 1963;119:945-951.
4. Vaillant GE: Prospective prediction of schizophrenic remission. *Arch Gen Psychiatry* 1964;11:509-518.
5. McCabe MS, Fowler RC, Cadoret RJ, et al: Symptom differences in schizophrenia with good and poor prognosis. *Am J Psychiatry* 1972;128:1239-1243.
6. Brockington IF, Wainwright S, Kendell RE: Manic patients with schizophrenic or paranoid symptoms. *Psychol Med* 1980;10:73-83.
7. American Psychiatric Association, Committee on Nomenclature and Statistics: *Diagnostic and Statistical Manual of Mental Disorders*, ed 3. Washington, DC, American Psychiatric Association, 1980.
8. Feighner JP, Robins E, Guze SB, et al: Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry* 1972;26:57-63.
9. Bland RC, Orn H: Schizophrenia: Diagnostic criteria and outcome. *Br J Psychiatry* 1979;134:34-38.
10. Brockington IF, Kendell RE, Leff JP: Definitions of schizophrenia: Concordance and prediction of outcome. *Psychol Med* 1978;8:387-398.
11. Spitzer R, Endicott J, Robins E: *Research Diagnostic Criteria*, Instrument No. 58. New York, New York State Psychiatric Institute, 1975.
12. Wing JK, Cooper JE, Sartorius N: *The Measurement and Classification of Psychiatric Symptoms*. London, Cambridge University Press, 1974.
13. Langfeldt G: Diagnosis and prognosis in schizophrenia. *Proc R Soc Med* 1960;53:1047-1052.
14. Carpenter WT, Strauss JS, Bartko JJ: Flexible system for the diagnosis of schizophrenia: Report from the WHO Pilot Study of Schizophrenia. *Science* 1973;182:1275-1278.
15. Schneider K: *Clinical Psychopathology*, Hamilton MW (trans). London, Grune & Stratton Inc, 1959, pp 133-134.
16. Cooper JE, Kendell RE, Gurland BJ, et al: *Psychiatric Diagnosis in New York and London*. New York, Oxford University Press, 1972.
17. Astrachan BM, Harrow M, Adler D, et al: A checklist for the diagnosis of schizophrenia. *Br J Psychiatry* 1972;121:529-539.
18. Kendell RE, Brockington IF, Leff JP: Prognostic implications of six alternative definitions of schizophrenia. *Arch Gen Psychiatry* 1979;36:25-31.
19. Hyler SE: Diagnostic reliability: Live versus case summary. Read before the annual meeting of the American Psychiatric Association, San Francisco, May 6, 1980.
20. Carey G, Gottesman II: Reliability and validity in binary ratings: Areas of common misunderstanding in diagnosis and symptom ratings. *Arch Gen Psychiatry* 1978;35:1454-1459.
21. Helzer JE, Clayton PJ, Pambakian R, et al: Reliability of psychiatric diagnosis: II. The test/retest reliability of diagnostic classification. *Arch Gen Psychiatry* 1977;34:136-141.
22. Spitzer RL, Endicott J, Robins E: Research Diagnostic Criteria: Rationale and reliability. *Arch Gen Psychiatry* 1978;35:773-782.
23. Goldberg DT: *The Detection of Psychiatric Illness by Questionnaire*, No. 21. Maudsley Monograph Series, New York, Oxford University Press, 1972.
24. Helzer JE, Clayton PJ, Pambakian R, et al: Concurrent diagnostic validity of a structured psychiatric interview. *Arch Gen Psychiatry* 1978;35:849-853.
25. Robins E, Guze SB: Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *Am J Psychiatry* 1970;126:983-987.
26. Strauss JS, Carpenter WT Jr: Prediction of outcome in schizophrenia: Five-year outcome and its predictors. *Arch Gen Psychiatry* 1977;34:159-163.
27. Kendell RE: The stability of psychiatric diagnoses. *Br J Psychiatry* 1974;124:352-356.
28. Morrison J, Clancy J, Crowe R, et al: The Iowa 500: I. Diagnostic validity in mania, depression, and schizophrenia. *Arch Gen Psychiatry* 1972;27:457-461.
29. Morrison J, Winokur G, Crowe R, et al: The Iowa 500: The first follow-up. *Arch Gen Psychiatry* 1973;29:678-682.
30. Tsuang MT, Woolson RF, Fleming JA: Long-term outcome of major psychoses: I. Schizophrenia and affective disorders compared with psychiatrically symptom-free surgical conditions. *Arch Gen Psychiatry* 1979;36:1295-1301.
31. Tsuang MT, Dempsey GM: Long-term outcome of major psychoses: II. Schizoaffective disorder compared with schizophrenia, affective disorders, and surgical control group. *Arch Gen Psychiatry* 1979;36:1302-1304.
32. Sartorius N, Jablensky A, Shapiro R: Two-year follow-up of the patients included in the WHO International Pilot Study of Schizophrenia. *Psychol Med* 1977;7:529-541.
33. Strauss JS, Carpenter WT Jr: Characteristic symptoms and outcome in schizophrenia. *Arch Gen Psychiatry* 1974;30:429-434.
34. Spitzer RL, Williams JBW: Classification of mental disorders and DSM-III, in Kaplan HI, Freedman AM, Sadock BJ (eds): *Comprehensive Textbook of Psychiatry*, ed 3. Baltimore, Williams & Wilkins Co, 1980, p 1040.
35. Shields J, Gottesman II: Cross-national diagnosis of schizophrenia in twins: The heritability and specificity of schizophrenia. *Arch Gen Psychiatry* 1972;27:725-730.